A Consistent Nonparametric Test of the Convexity of Regression Based on Least Squares Splines [1]

by

Cheikh A.T. DIACK

Laboratoire de Statistique et Probabilités UMR CNRS C5583 Université Paul Sabatier 118 route de Narbonne 31062 Toulouse cedex 4, France.

e-mail:diack@cict.fr

Abstract

This paper provides a test of convexity of a regression function. This test is based on the least squares splines. The test statistic is shown to be asymptotically of size equal to the nominal level, while diverging to infinity if the convexity is misspecified. Therefore, the test is consistent against all deviations from the null hypothesis.

1 INTRODUCTION

Tests of convexity of a regression function is one of the most important problems in econometrics. Indeed, "*The General Theory of Employment, Interest, and Money* emphasized the central importance of the consumption function and explicitly argued that the consumption function is concave" (Carroll & Kimball 1996). Economic theory predicts also the convexity of functions like for example Bernoulli utility function, cost function, production function, Engels curves, ... Otherwise, the Human Capital theory argued that the relationships between the logarithm of wage and the experience is concave.
On the other hand, psychologists have worried for over a century about whether subjective reports about physical magnitudes like length, weigth, area, luminance and etc. have a convex or concave relationships to corresponding measurement. Also, this convexity problem is very closely connected to the order-restricted hypothesis testing problems describe in references such as Robertson and al. (1988).

There are some papers in the statistics literature dealing with nonparametric hypothesis testing convexity of the regression function. The work along this line includes Schlee (1980), Yatchew (1992), Diack (1996), Diack (1997), Diack & Thomas (1998) and Diack (1998).
Schlee (1980) in a nonparametric regression model with random design used an estimator of the second derivative of the regression function. His test statistic requires computing the distribution of the supremum of this normalized estimator

[1] *key words:least squares estimator, test of convexity, Likelihood ratio test, convex cone.*

over an intervall. But this method imposes some theorical difficulties. To overcome the problem, he proposes a sequence of points from the interval and uses the theory of maximal deviation to obtain the distribution of the test statistic under the null hypothesis. However, this work no discuss asymptotic results or pratical implementation.
Yatchew's test (with semi-parametric model) is based on comparing the nonparametric sum of squared residuals with convexity contraints, with the nonparametric sum of squared residuals without contraints. Yatchew's approach relies on sample splitting which products a lost of efficiency. He gives a heuristic proof of the consistency of the test. Diack (1998) adapt respectively Schlee's idea and Yatchew's idea in a nonparametric model with fixed design to construct two other tests of convexity for which he gives new asymptotic results of convergence.
Diack and Thomas (1998) use least squares splines estimator and develope in nonparametric model with deterministic design, a non-convexity test which is consistent for some alternative hypothesis. A small simulation study in Diack(1996) shows that the finite sample of this test is quite satisfactory.

In this paper, we propose a new test of convexity of a regression function in a nonparametric model. Our test uses, as Diack and Thomas's test, a cubic spline estimator which allows us to formulate the convexity hypothesis in a very simple way. Hence, our problem becomes roughly, a problem to test a multivariate normal mean with composite hypotheses determined by linear inequalities.

The remainder of this paper is organized as follows. In section 2, we introduce the nonparametric regression model and the hypotheses to be tested. After, we recall some properties of the cubic spline estimator. Section 3 describes our test of convexity of the regression function and section 4 is devoted to a discussion and demonstration of some properties of the test.

2 PRELIMINARIES

2.1 The model and the hypotheses

Consider the nonparametric regression model:

$$y_{ij} = f(x_i) + \varepsilon_{ij}, i = 1, ..., r, j = 1, ..., n_i \quad \text{with} \quad x_i \in (0,1), i = 1, ..., r.$$

At each deterministic design point $x_i, (i = 1, ..., r), n_i$ measurements are taken. The probability measure assigning mass $\mu_i = n_i/n$ to the point $x_i(\sum \mu_i = 1)$ is referred to as the design and will be denoted by μ^n. We assume that the random errors ε_{ij} are uncorrelated and identically distributed with mean zero. Their variance σ^2 will be assumed unknown. Finally f is an unknown smooth regression function.

In what follows, we will assume some regularity conditions on f.

The following class of functions were use by Diack and Thomas (1998) to construct a test of non-convexity.

For $l \in I\!N$ and $M, L > 0$, let

$$\mathcal{F}_{l,M} = \{f \in \mathcal{C}^{l+1}(0,1) : \sup_{0 \le x \le 1} \mid f^{((l+1)\wedge 4)}(x) \mid \le M\}.$$

We intend to contruct a test of convexity of the regression function f. Thus the null hypothesis to be tested is that the convexity of f is correct:

H_0 : "f is convex"

while, without a specific alternative hypothesis, the alternative to be tested will be that the null is false:

H_1 : "f is non-convex."

Thus the alternative encompasses all the possible departure from the null hypothesis.

In what follows, a testing problem with null hypothesis H_0 and alternative H_1 is denoted by $[H_0, H_1]$.

We will use a cubic spline estimator and characterizing convexity in the set of collection of all polynomial cubic splines, so we transform our problem at a test of a multivariate normal mean with composite hypotheses determined by linear inequalities.

2.2 The Cubic Spline Estimator

Let p be a positive continuous density on (0,1). We assume that

$$\min_{0 \le x \le 1} p(x) > 0.$$

Let $\eta_0 = 0 < \eta_1 < ... < \eta_{k+1} = 1$ be a subdivision of the interval (0,1) by k distincts points defined by

$$\int_0^{\eta_i} p(x)\, dx = i/(k+1), \quad i = 0, ..., k+1. \tag{2.20}$$

Let $\delta_k = \max_{0 \le i \le k}(\eta_{i+1} - \eta_i)$, we see that

$$\delta_k / \min_i(\eta_{i+1} - \eta_i) \le \max_x p(x) / \min_x p(x). \tag{2.21}$$

For each fixed set of knots of the form (2.20), we define $\mathcal{S}$(k,d) as the collection of all polynomial splines of order d (degree$\leq$ d-1) having for knots $\eta_1 < ... < \eta_k$. The class $\mathcal{S}$(k,d) of such splines is a linear space of functions of dimension (k+d). A basis for this linear space is provided by the B-splines(see Schumaker (1981)). Let $\{N_1, .., N_{k+d}\}$ the set of normalized B-splines associated to the following non-decreasing sequence $\{t_1, .., t_{k+2d}\}$:

$$\begin{cases} t_1 \leq t_2 \leq ... \leq t_d = 0 \\ t_{2d+k} \geq t_{2d+k-1} \geq ... \geq t_{d+k+1} = 1 \\ t_{d+l} = \eta_l \quad \text{for} \quad l = 1, .., k \end{cases}$$

The reader is referred to Schumaker (1981) for a discussion of these B-splines.

In what follows, we shall only work with the class of cubic splines: $\mathcal{S}$(k,4). It will be convenient to introduce the following notations:

$$N(x) = (N_1(x), .., N_{k+4}(x))' \in \mathbb{R}^{k+4}.$$

$$F = (N(x_1), .., N(x_r)); (k+4) \times r \quad \text{matrix}.$$

We will denote by $\widehat{f}_n$ the least squares spline estimator of f :

$$\widehat{f}_n(x) = \sum_{p=1}^{k+4} \widehat{\theta}_p N_p(x) \tag{2.22}$$

where

$$\widehat{\Theta} = (\hat{\theta}_1, ..., \hat{\theta}_{k+4})' = \arg \min_{\Theta \in \mathbb{R}^{k+4}} \sum_{i=1}^{r} \sum_{j=1}^{n_i} (y_{ij} - \sum_{p=1}^{k+4} \theta_p N_p(x_i))^2. \tag{2.23}$$

Let

$$\bar{y_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}, \bar{\varepsilon}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \varepsilon_{ij},$$

$$\bar{Y} = (\bar{y_1}, .., \bar{y_r})', \bar{\varepsilon} = (\bar{\varepsilon_1}, .., \bar{\varepsilon_r})', f_\Delta = (f(x_1), .., f(x_r))'$$

Let $\mathcal{D}(\mu^n)$ be the $r \times r$ diagonal matrix with diagonal elements $\mu_1, ..., \mu_r$, then, basic least squares arguments prove that:

$$\widehat{\Theta} = M^{-1}(\mu^n) F \mathcal{D}(\mu^n) \bar{Y} \quad \text{with} \quad M(\mu^n) = \sum_{i=1}^{r} N(x_i) N'(x_i) \mu_i = F \mathcal{D}(\mu^n) F'.$$

Asymptotic properties of this estimator have been established in Argarwal and Studden(1980).

Remark that the first moment of $\widehat{f}_n$ is given by

$$\mathbb{E}\widehat{f}_n(x) = N(x)' M^{-1}(\mu^n) F \mathcal{D}(\mu^n) f_\Delta.$$

Thus, if f is a cubic spline function (that is to say there is Θ such that $f_\Delta = F'\Theta$) then $\widehat{f}_n$ is unbiased and $\mathbb{E}\widehat{\Theta} = \Theta$. We will use below this property to construct our test.

3 TEST STATISTIC

3.1 Convexity in $\mathcal{S}$(k,4)

First of all, characterize convexity in the class $\mathcal{S}$(k,4). Remark that, if a function g is a cubic spline then, its second derivative is a linear function between any pair of adjacent knots η_i and η_{i+1}, and it follows that g is a convex function in the interval $\eta_i \leq x \leq \eta_{i+1}$ if and only if $g''(\eta_i)$ and $g''(\eta_{i+1})$ are both non negative (this property was used by Dierckx (1980) to define a convex estimator).
For a function g in the class $\mathcal{S}$(k,4), we can write:

$$g(x) = \sum_{p=1}^{k+4} \theta_p N_p(x) \quad \text{with} \quad \Theta = (\theta_1, ..., \theta_{k+4})' \in I\!R^{k+4}$$

$$\text{Then:} \qquad g''(\eta_l) = \sum_{p=1}^{k+4} \theta_p N''_p(\eta_l) = \sum_{p=1}^{k+4} \theta_p d_{p,l},$$

where the coefficients $d_{p,l}$ are easily calculated from the knots (see Dierckx (1980))

$$\left\{ \begin{array}{l} d_{p,l} = 0 \quad \text{if} \quad p \leq l \quad \text{or} \quad p \geq l+4 \\ \qquad d_{l+1,l} = \frac{6}{(t_{l+5}-t_{l+2})(t_{l+5}-t_{l+3})} \\ \qquad d_{l+3,l} = \frac{6}{(t_{l+6}-t_{l+3})(t_{l+5}-t_{l+3})} \\ \qquad d_{l+2,l} = -(d_{l+3,l} + d_{l+1,l}) \end{array} \right. \quad \text{for} \quad l = 0, ..., k+1$$

Let $b_l = (0, 0, .., 0, -d_{l+1,l}, -d_{l+2,l}, -d_{l+3,l}, 0, .., 0)' \in I\!R^{k+4} and \quad \Theta = (\theta_1, .., \theta_{k+4})'$, then

$$g''(\eta_l) = -b'_l\Theta.$$

Hence, we see that a cubic spline g is a convex function if and only if $b'_l\Theta \leq 0$ for all l=0,...,k+1.
Hence, the idea of our test follows from this property. Indeed, we have already mentionned in the section 2 that, if f is a cubic spline then

$$I\!E\widehat{\Theta} = \Theta.$$

Therefore to build a test of convexity in the case where the regression function f is a cubic spline function, is the same to build a test for the hypotheses concerning linear inequalities and random vector mean. More precisely, the test $[H_0, H_1]$ is equivalent in this case at the following test with null hypothesis

$H_0' : b'_l\Theta \leq 0$ for all $l = 0, ..., k+1$.

against the alternative

$H_1' : \exists l : b'_l\Theta > 0$
where Θ is the mean of the random vector $\widehat{\Theta}$.

On the other hand, Beatson (1982) shows that for a smooth and convex function $f \in \mathcal{C}^m(0,1)(0 \leq m \leq 3)$, the uniform distance between f and the set $\mathcal{S}^{**}$(k,4) of convex functions of $\mathcal{S}$(k,4) tends to zero when the mesh size δ_k tends to zero (see lemma 3 below).

A testing problem in the form $[H_0^{'}, H_1^{'}]$ is related to the one-sided testing problem in multivariate analysis and has been studied by several authors (Bartholomew 1961, Kudô 1963, Nüesch 1966, Kudô and Choi 1975, Shapiro 1985 and more recently by Raubertas and al. 1986 and Robertson and al. 1988).

3.2 One-sided Test

Let Y be a random vector distributed as $\mathcal{N}_q(\Theta, \Sigma_q)$ $(q \in I\!N, \quad q > 0)$ where Σ_q is a known nonsingular matrix.
We consider the following testing problem:

$$\text{null hypothesis, } H_0^{'} : b'_l\Theta \leq 0 \quad (l = 0, ..., k+1).$$

$$\text{alternative, } H_1^{'} : \exists l : b'_l\Theta > 0.$$

In this paper we identify an hypothesis with the corresponding set of parameters. For example, we write $H_0^{'} = \{\Theta \in I\!R^q : b'_l\Theta \leq 0\}$.
The likelihood function is $L = c_0 exp(-\frac{1}{2} \mid\mid Y - \Theta \mid\mid^2_{\Sigma_q^{-1}})$ where c_0 is a positive constant independent of Θ. Thus, the likelihood ratio for the problem $[H_0^{'}, H_1^{'}]$ is given by:

$$\lambda = \frac{\sup_{H_o^{'}} L}{\sup_{H_o^{'} \cup H_1^{'}} L} = exp(-\frac{1}{2} \inf_{x \in H_o^{'}} \mid\mid Y - x \mid\mid^2_{\Sigma_q^{-1}}).$$

So, to determine the test statistic under the null hypothesis, we need to resolve the following nonlinear programming: $\inf_{x \in H_o^{'}} \mid\mid Y - x \mid\mid^2_{\Sigma_q^{-1}}$.
Remark that $H_o^{'}$ is a polyhedral convex cone. Hence, for a given Y, this infinimum is attained at unique point denoted by $\Pi_{H_o^{'}}(Y)$ and represents the square distance from Y to $H_o^{'}$.
Thus, the likelihood ratio test (LRT) rejects $H_1^{'}$ for large values of

$$\bar{\chi}^2 = \inf_{x \in H_o^{'}} \mid\mid Y - x \mid\mid^2_{\Sigma_q^{-1}} .$$

Shapiro (1985) showed, in a study of the distribution of a minimum discrepancy statistic, that if $H_o^{'}$ is a any convex cone and if $\Theta = 0$, then the distribution of $\bar{\chi}^2$ statistic, called chi-bar-squared statistics, is a mixture of chi-squared distributions. Raubertas and al.(1986) generalize the one-side testing problem to allow hypotheses involving homogeneous linear inequality restrictions. This framework includes the hypotheses of monotonicity, nonnegativity, and convexity. Here we give an immediate consequence of theorem 3.1 of Shapiro (1985). For that purpose we shall use some geometrical properties of polyhedral cones.

3.2.1 Polyhedral convex cones

Let $\{a_1, ..., a_p\}$ be a set of vectors in $I\!R^q$ (with $p \leq q$) and let $\mathcal{C}$ be a convex polyhedral cone defined by $\{a_1, ..., a_p\}$ that is,

$$\mathcal{C} = \{x \in I\!R^q \ : x = \sum_{i=1}^{p} \lambda_i a_i, \quad \lambda_i \geq 0, \quad i = 1, ..., p\}.$$

The polar cone $\mathcal{C}^\circ$ of a cone $\mathcal{C}$ is defined by

$$\mathcal{C}^\circ = \{x \in I\!R^q \ : x'y \leq 0, \quad \forall y \in \mathcal{C}\}.$$

It is easy to see that

$$\mathcal{C}^\circ = \{x \in I\!R^q \ : a_i^{'} x \leq 0, \quad i = 1, ..., p\}.$$

Note that $\mathcal{C}^\circ$ is also a polyhedral convex cone and $(\mathcal{C}^\circ)^\circ = C$.

The boundary of $\mathcal{C}^\circ$ is given by

$$\partial\mathcal{C}^\circ = \{x \in I\!R^q \ : a_i^{'} x \quad {}_* \leq 0, i = 1, ..., p\},$$

where the notation ${}_* \leq$ mean (as in Sasabuchi 1980) that at least one equality hold in this sequence of inequalities.

The polar cone $\mathcal{C}^\circ$ has p faces F_j;

$$F_j = \{x \in I\!R^q \ : a_j^{'} x = 0, a_i^{'} x \leq 0, i = 1, ..., p\}, j = 1, ..., p.$$

The dimension of each face is at most $p - 1$.
For the definition and basic properties of faces and polar cone the reader is reffered to Rockafeller (1970).
To a face F_j, we denote by $I\!P_{F_j}$ the symmetric idempotent matrix giving the orthogonal projection onto the space generated by F_j. As above, for all x, $\Pi_{\mathcal{C}^\circ}$ represents the projection of x onto $\mathcal{C}^\circ$.
Now, we shall need the following result which in various forms has been used by several authors (Kudô (1963), Shapiro (1985)).

Lemma 1 *For all $x \in I\!R^q$, we have:*
$\Pi_{\mathcal{C}^\circ}(x) \in F_j$ *if and only if* $P_{F_j}(x) \in \mathcal{C}^\circ$ *and* $x - P_{F_j}(x) \in \mathcal{C}$ *and in this case:*

$$\Pi_{\mathcal{C}^\circ}(x) = P_{F_j}(x) \quad and \quad x - P_{F_j}(x) = \frac{a'_j x}{\|\ a_j \ \|^2} a_j.$$

In some pratical situations there may be all the face F_j are exactly $p-1$ dimensional. That may be depends on the configuration of the vectors $a_1, ..., a_p$. For these cases, we prepare some definitions and theorems which are a restatement of definition 2.1, 2.2 and theorem 2.1 of Sasabuchi (1980) in this setting.

Definiton 1 A vector a_j is said to be redundant in $\{a_1, ..., a_p\}$ if

$$\{x : a_i^{'}x \geq 0, i = 1, ..., p\} = \{x : a_i^{'}x \geq 0, i = 1, ..., p, i \neq j\}.$$

or equivalently, there does not exist a vector x such that $a_i^{'}x \geq 0 (i \neq j), a_j^{'}x < 0$.

Definition 2 A set of vectors $\{a_1, ..., a_p\}$ is said to be with positive relations if there exist nonnegative numbers $\lambda_1, ..., \lambda_p$, not all simultaneously zero, such that $\sum_i \lambda_i a_i = 0$; otherwise the set is said to be without positive relations.

Remark 1 If $\{a_1, ..., a_p\}$ is with positive relations, it is linearly dependent; equivalently, if $\{a_1, ..., a_p\}$ is linearly independent, it is without positive relations.

Theorem 1 *Suppose that $\{a_1, ..., a_p\}$ is without positive relations and has no redundant vector in it. Then all the faces $F_j (j = 1, ..., p)$ are exactly $p-1$ dimensional.*

In our case, $\mathcal{C}^{\circ} = H_{\circ}^{'}$ with $p = k+2$ and it is easy to see that $\{b_0, ..., b_{k+1}\}$is without positive relations and has no redundant vector in it.

3.2.2 The distribution of $\bar{\chi}^2$

The following result is an immediate consequence of theorem 3.1 of Shapiro (1985) and theorem 1 above.

Theorem 2 *Let Y be a random vector distributed as $\mathcal{N}_q(\Theta, \Sigma_q)$ where $\Theta \in I\!R^q$ and Σ_q is a known nonsingular matrix. Suppose that $\{a_1, ..., a_p\}$ is without positive relations and has no redundant vector in it. If $\Theta = 0$, then the random variable $\bar{\chi}^2$ is distributed as a mixture of chi-squared distributions, namely*

$$P(\bar{\chi}^2 \geq s^2) = \omega_0 P(\chi_0^2 \geq s^2) + (\sum_{j=1}^{p} \omega_j) P(\chi_1^2 \geq s^2) + \omega_q P(\chi_q^2 \geq s^2) \quad (3.20)$$

with

$$\omega_0 = P(Y \in H_{\circ}^{'}) = P({a'}_i Y \leq 0, \quad i = 1, ..., p),$$

$$\omega_j = P({a'}_j Y \leq 0) P({a'}_i Y - \frac{{a'}_i \Sigma_q a_j}{({a'}_j \Sigma_q a_j)} {a'}_j Y \leq 0, \quad i = 1, ..., p), \quad j = 1, ..., p$$

and

$$\omega_q = P(Y \in (H_{\circ}^{'})^{\circ}).$$

Moreover

$$\sum_{j=0}^{p} \omega_j + \omega_q = 1.$$

See Diack (1997) for a proof of this theorem. The result of theorem 2 shows that the distribution of $\bar{\chi}^2$ when $\Theta = 0$, is a mixture of chi-squared distributions. So, to calculate the probabilities in the right-hand side of (3.20), the values of $\omega_j, (j = 0, ..., p)$ are needed. However, even for moderate $q(q > 3)$, good closed form expressions for these level probabilities have not found. Thus approximations are of interest. For this, one can use Monte Carlo method (see Diack (1997)).
Note that the coefficients ω_j depend of the vectors a_j and the matrix Σ_q. Hence, in what follows, we denoted $\bar{\chi}^2$ by $\bar{\chi}^2_{\Sigma_q}(p)$.
Questions concerning the determination of the distribution of $\bar{\chi}^2_{\Sigma_q}(p)$ for any point of null hypothesis are unresolved. However, the following lemma which can find in Robertson and al. (1988), prove that, theorem 2 suffices to construct a test of size α.

Lemma 2 *Let $\mathcal{K}$ be a closed convex cone. For x in $I\!R^q$, $x_{\mathcal{K}}$ denotes its projection onto $\mathcal{K}$ with respect any inner product $< .,. >$ and $\|\ .\ \|$ the associated norm. If $z \in \mathcal{K}$, then*

$$\| x + z - (x + z)_{\mathcal{K}} \| \leq \| x - x_{\mathcal{K}} \| .$$

Remark 2 Recall that $Y \rightsquigarrow \mathcal{N}_q(\Theta, \Sigma_q)$. Let P_{Θ_0} be the law of probabilities of Y under the hypothesis $\Theta = \Theta_0$. Then, for all $\Theta_0 \in \mathcal{K}$, we have:
under the hypothesis $(\Theta = \Theta_0)$, $\quad Y \rightsquigarrow \mathcal{N}_q(\Theta_0, \Sigma_q)$,
under the hypothesis $(\Theta = 0)$, $\quad Y + \Theta_0 \rightsquigarrow \mathcal{N}_q(\Theta_0, \Sigma_q)$,
hence, under the hypothesis $(\Theta = \Theta_0)$ the law of $\| Y - Y_{\mathcal{K}} \|^2_{\Sigma_q^{-1}}$ is the same that the law of $\| Y + \Theta_0 - (Y + \Theta_0)_{\mathcal{K}} \|^2_{\Sigma_q^{-1}}$ under the hypothesis $(\Theta = 0)$.
Now, using lemma 1, we see that

$$P_{\Theta_0}(\| Y - Y_{\mathcal{K}} \|^2_{\Sigma_q^{-1}} \geq s) = P_0(\| Y + \Theta_0 - (Y + \Theta_0)_{\mathcal{K}} \|^2_{\Sigma_q^{-1}} \geq s)$$

$$\leq P_0(\| Y - Y_{\mathcal{K}} \|^2_{\Sigma_q^{-1}} \geq s).$$

The last inequality mean that, if $\mathcal{K}$ is a set of parameters corresponding to null hypothesis of a test statistic given by $\| Y - Y_{\mathcal{K}} \|^2_{\Sigma_q^{-1}}$, then the maximal level of this test is obtain for $\Theta = 0$. In this case, $(\Theta = 0)$ is so-called the least sub-hypothesis of $\mathcal{K}$.
Lemma 2 has the following consequence: the size-α likelihood ratio test with null hypothesis $H_\circ^{'}$ versus the alternative hypothesis $H_1^{'}$ is the test with reject the null hypothesis if

$$\bar{\chi}^2_{\Sigma_q}(p) \geq s^2_{\alpha,p}$$

where $s^2_{\alpha,p}$ is defined by

$$(\sum_{j=1}^{p}\omega_j)P(\chi_1^2 \geq s^2_{\alpha,p}) + (1-\omega_0-(\sum_{j=1}^{p}\omega_j))P(\chi_q^2 \geq s^2_{\alpha,p}) = \alpha. \quad (3.21)$$

Hence $s^2_{\alpha,p}$ is a function of the weights ω_j. Now, the following result give a sufficient condition of convergence of the power of the test.

Theorem 3 *Under the assumptions of theorem 2, the power function of* $\bar{\chi}^2_{\Sigma_q}(p)$ *converges uniformly to one as* $\min_{x\in H_\circ'} \| \Theta - x \|^2_{\Sigma_q^{-1}} \longrightarrow +\infty$.

Proof
Let T be a $q\times q$ nonsingular matrix such that $T\Sigma_q T' = I_q$, that is $\Sigma_q = T^{-1}(T^{-1})'$, and make the transformation

$$X = TY, \quad U = T\Theta.$$

Then X is a random vector distributed as $\mathcal{N}_q(U, I_q)$. Define the set of vectors $\{c_1, ..., c_p\}$ as

$$c'_j = a'_j T^{-1}, \quad (j = 1, ..., p).$$

It is esay to see that the set of vectors $\{c_1, ..., c_p\}$ is without positive relations and has no redundant vector in it.
We have $a'_j\Theta = c'_jU \quad (j = 1, ..., p)$, and hence the problem $[H_\circ', H_1']$ is transformed to the following problem $[H_\circ'', H_1'']$:
$H_\circ'' : c'_jU \leq 0 \quad (j = 1, ..., p)$
$H_1'' : \exists j : c'_jU > 0$.
We can write

$$\bar{\chi}^2_{\Sigma_q}(p) = \min_{x\in H_\circ''} \| X - x \|^2 = \| X - \Pi_{H_\circ''}(X) \|^2 .$$

On the other hand,

$$\| U - \Pi_{H_\circ''}(U) \| \leq \| U - \Pi_{H_\circ''}(X) \| \leq \| X - U \| + \| X - \Pi_{H_\circ''}(X) \| .$$

hence, the result follows from the assumption

$$\| U - \Pi_{H_\circ''}(U) \| \longrightarrow +\infty.\Box$$

The test statistic requires computing the projection $\Pi_{H_\circ'}(Y)$ of Y. However, a good closed-form solution has not found. Hence, this problem requires extensive numerical work to obtain solution. We propose an algorithm based on succesive projections which has been introduced by Dykstra (1983) (see also Boyle and Dykstra (1985)). This algorithm determines the projection of a point X of any real Hilbert space onto the intersection $\mathcal{K}$ of convex set $\mathcal{K}_j \quad (j = 1, ..., p)$ and

it is meant for applications where projections onto the $\mathcal{K}_j$'s can be calculated relatively easily. Let $\mathcal{K}$ be a closed convex cone in $I\!R^q$. We suppose that, $\mathcal{K}$ can be written as $\bigcap_{j=1}^{p} \mathcal{K}_j$ and each $\mathcal{K}_j$ is also closed convex cone. For all $X \in I\!R^q$, we denote by $X_{\mathcal{K}}^{\Gamma}$ the $\Gamma-$ projection onto $\mathcal{K}$, where Γ is a positive definite matrix. The algorithm consists of repeated cycles and every cycle contains p stages.
Let X_{mi}^{Γ} be the appproximation of $X_{\mathcal{K}}^{\Gamma}$ given by Dykstra's algorithm at the ith stage of mth cycle.
The following result (see Boyle & Dykstra (1985)) proves that the algorithm converges correctly.

Theorem 4 *For any* $(1 \leq i \leq p)$, *the sequence* $\{X_{mi}^{\Gamma}\}$ *converges strongly to* $X_{\mathcal{K}}^{\Gamma}$ *i.e.* $\| X_{mi}^{\Gamma} - X_{\mathcal{K}}^{\Gamma} \|_{\Gamma} \to 0$ *as* $m \to +\infty$.

Application: Let $\mathcal{K} = H_{\circ}^{'}$ be the null hypothesis of the problem $[H_{\circ}^{'}, H_1^{'}]$ and let $\mathcal{K}_i = \{x \in I\!R^q \ : b_i^{'} x \leq 0\}$. Let $\Gamma = \Sigma_q^{-1}$ be the covariance matrix of Y. For all $m \in I\!N, \quad m > 0$, we defined $\bar{\chi}^2_{\Sigma_q}(p, m)$ by

$$\bar{\chi}^2_{\Sigma_q}(p, m) = \| Y - Y_{mp}^{\Sigma_q^{-1}} \|^2_{\Sigma_q^{-1}}$$

where $Y_{mp}^{\Sigma_q^{-1}}$ is given by the pth stage of mth cycle of the Dykstra's algorithm. We have then the following equality:

$$\bar{\chi}^2_{\Sigma_q}(p) = \| Y - Y_{H_{\circ}^{'}}^{\Sigma_q^{-1}} \|^2_{\Sigma_q^{-1}} = \| Y - Y_{mp}^{\Sigma_q^{-1}} \|^2_{\Sigma_q^{-1}} + \| Y_{mp}^{\Sigma_q^{-1}} - Y_{H_{\circ}^{'}}^{\Sigma_q^{-1}} \|^2_{\Sigma_q^{-1}}$$

$$+2 < Y - Y_{mp}^{\Sigma_q^{-1}}, Y_{mp}^{\Sigma_q^{-1}} - Y_{H_{\circ}^{'}}^{\Sigma_q^{-1}} >_{\Sigma_q^{-1}}$$

where $<,>_{\Sigma_q^{-1}}$ is the inner product on $I\!R^q$ defined by Σ_q^{-1}. Using now theorem 4, we see that

$$\| Y_{mp}^{\Sigma_q^{-1}} - Y_{H_{\circ}^{'}}^{\Sigma_q^{-1}} \|^2_{\Sigma_q^{-1}} \to 0 \qquad \text{a.s. as} \qquad m \to +\infty.$$

In the same way, it can be shown that for fixed p and q

$$< Y - Y_{mp}^{\Sigma_q^{-1}}, Y_{mp}^{\Sigma_q^{-1}} - Y_{H_{\circ}^{'}}^{\Sigma_q^{-1}} >_{\Sigma_q^{-1}} \to 0 \qquad \text{a.s. as} \qquad m \to +\infty.$$

Hence, $\bar{\chi}^2_{\Sigma_q}(p, m)$ converges almost surely to $\bar{\chi}^2_{\Sigma_q}(p)$ as m tends to infinity. Therefore, to implement the test statistic, we will used $\bar{\chi}^2_{\Sigma_q}(p, m)$ instead $\bar{\chi}^2_{\Sigma_q}(p)$.
We can now define our convexity test.

3.3 Definition of the test

Consider the problem $[H_{\circ}, H_1]$ where $H_{\circ}$ means the regression function f, is convex and H_1 is the unrestricted alternative to be tested.

Let $\widehat{\Theta}$ be the solution of the quadratic programming problem (2.23).
Let define $\widehat{\Theta}_{m,k_n+2}$ by

$$\widehat{\Theta}_{m,k_n+2} = \widehat{\Theta}^{\Sigma_n}_{m,k_n+2}$$

with $\Sigma_n^{-1} = \frac{\sigma^2}{n} M^{-1}(\mu^n)$ and where $\widehat{\Theta}^{\Sigma_n}_{m,k_n+2}$ given by the $(k_n+2)^{th}$ stage of the m^{th} cycle of Dykstra's algorithm.
Like this, we will defined our test of convexity by rejecting $H_\circ$ when

$$\bar{\chi}^2_{\frac{\sigma^2}{n}M^{-1}(\mu^n)}(k_n+2, m) = \frac{n}{\sigma^2} \mid\mid \widehat{\Theta} - \widehat{\Theta}_{m,k_n+2} \mid\mid^2_{M(\mu^n)} \geq s^2_{\alpha,k_n+2}, \tag{3.31}$$

where s^2_{α,k_n+2} is defined by (3.21).

4 ASYMPTOTIC PROPERTIES

Note that the test procedure requires the knowledge of the variance σ^2. However, the following results can be considered approximately true if σ is unknown but, in order to compute the statistic, we need a consistent estimate of σ. This can be obtained in the case of the least squares estimator, using $\widehat{\sigma}_n = \frac{1}{n-(k+4)} \sum_{i=1}^{r} (\bar{y}_i - \widehat{f}_n(x_i))^2$ or alternatively, any consistent estimator based on nonparametric regression techniques.
In what follows, we assume that μ^n converges to a design measure μ, where μ is an absolutely continuous measure. We denote by H_n and H the cumulative distribution function of μ^n and μ respectively. The number of knots will be function of the sample size $k = k_n$. The critical region of the test is

$$\Lambda_{n,m} = \{\bar{\chi}^2_{\frac{\sigma^2}{n}M^{-1}(\mu^n)}(k_n+2, m) = \frac{n}{\hat{\sigma}^2_n} \mid\mid \widehat{\Theta} - \widehat{\Theta}_{m,k_n+1} \mid\mid^2_{\Sigma_n} \geq s^2_{\alpha,k_n+2}\}.$$

Now, we are ready for the main result of this section.

Theorem 5 *Let $f \in \mathcal{F}_{l,M}$ with $l \geq 3$. Let us consider the problem $[H_\circ, H_1]$. Then, under the following assumptions*

$$(i) \varepsilon_{ij} \rightsquigarrow \mathcal{N}(0, \sigma^2) \quad \textit{and iid},$$

$$(ii) \sup_{0 \leq x \leq 1} \mid H_n(x) - H(x) \mid = \circ({k_n}^{-1}), \qquad \textit{as} \quad k_n \longrightarrow +\infty$$

$$(iii) \lim_{n \longrightarrow +\infty} r^{1/2} n^{1/2} {\delta_{k_n}}^3 (\sup_{1 \leq i \leq r} \mu_i)^{1/2} = 0,$$

the test is asymptotically size α. More precisely,

$$\lim \sup_{n \longrightarrow +\infty} \sup_{f \in H_0} \lim_{m \longrightarrow +\infty} \mathcal{P}_f(\Lambda_{n,m}) = \alpha.$$

Moreover, if

$$(iv) \lim_{n \longrightarrow +\infty} n^{1/2} \delta^{5/2}_{k_n} = +\infty$$

then, the test statistic is diverging to infinity: it is consistent i.e

$$\lim_{n \longrightarrow +\infty} \lim_{m \longrightarrow +\infty} \mathcal{P}_f(\Lambda_{n,m}) = 1.$$

The proof of theorem 5 is given in Appendix.

Remarks

- Assumption (ii) is the same as in Aggarwal and Studden (1980)
- It is easy to see that assumption (iii) implies that

$$\lim_{n \longrightarrow +\infty} k_n = +\infty$$

- For a uniform design, i.e $\mu_i = 1/r, i = 1, ..., r$ then (iii) is equivalent to

$$\lim_{n \longrightarrow +\infty} n^{1/2} \delta_{k_n}{}^{(m+1)\wedge 4} = 0$$

  This is the case for example when there are n distinct points for n measurements(r=n).

Discussion: We have proposed a consistent test of convexity of a regression function in a nonparametric model. While it appears difficult to impose properties such as concavity on nonparametric local averaging estimators, this restriction is readily introduced by using cubic spline estimator. Hence, the idea of the test exploits the close connection between the convexity problem and the hypothesis testing problems concerning linear inequalites and normal means. A simulation study in Diack (1997) shows the usefulness of this method and that the finite sample behavior of the test is quite satisfactory.
It would be nice to extend this framework at the case where the errors are no necessary gaussian and to studie the behavior of the tests under the local alternatives.
A test of monotonicity can be readily constructed paralleling the above convexity test with quadratic splines instead of cubic splines. This additionnal step is still under study.

Appendix For the proof of the theorem 5, we need to use lemma 3 which is obtained by a straightforward manipulation of results of Schumaker (1981). It gives a sup norm error bound when approximating a smooth and convex function with a convex cubic spline (see also, Beatson (1982) and Diack (1997)).

Lemma 3 *Let $l \geq 3$. There is a constant c such that for all function $f \in \mathcal{F}_{l,M}$,there is S in $\mathcal{S}(k_n, 4)$ such that:*

$$\sup_{0 \leq x \leq 1} \mid f^{(j)}(x) - S^{(j)}(x) \mid \leq c\delta_{k_n}^{4-j}, \quad j = 0, ..., 3.$$

Moreover, if f is convex then S is also convex

Proof of theorem 5:
Let us recall that

$$\widehat{\Theta} = M^{-1}(\mu^n) F \mathcal{D}(\mu^n) \bar{Y}.$$

From Lemma 3, there is a function S in $\mathcal{S}(k_n, 4)$, such that :

$$\sup_{0 \le x \le 1} \mid f^{(j)}(x) - S^{(j)}(x) \mid \le c \delta_{k_n}^{4-j} \qquad \text{for all} \qquad j = 0, ..., 3.$$

$S \in \mathcal{S}(k_n, 4)$ hence, there exists $\Theta \in I\!\!R^{k_n+4}$ such that $S(x) = N'(x)\Theta$.
Let

$$S_\Delta = (S(x_1), .., S(x_r))' = (N'(x_1)\Theta, .., N'(x_r)\Theta)' = F'\Theta.$$

Then,

$$M^{-1}(\mu^n) F \mathcal{D}(\mu^n) S_\Delta = M^{-1}(\mu^n) F \mathcal{D}(\mu^n) F'\Theta = \Theta.$$

Let

$${}_s\widehat{\Theta} = M^{-1}(\mu^n) F \mathcal{D}(\mu^n)(S_\Delta + \bar{\varepsilon}).$$

Then

$$I\!\!E_s\widehat{\Theta} = \Theta \qquad \text{and} \qquad {}_s\widehat{\Theta} \rightsquigarrow \mathcal{N}(\Theta, \frac{\sigma^2}{n} M^{-1}(\mu^n)).$$

Therefore, we can write

$$\widehat{\Theta} = {}_s\widehat{\Theta} + B_n \qquad \text{and} \qquad I\!\!E\widehat{\Theta} = \Theta + B_n$$

with $B_n = M^{-1}(\mu^n) F \mathcal{D}(\mu^n)(f_\Delta - S_\Delta)$.
Now, let us recall that the test statistic is given by

$$\bar{\chi}^2_{\frac{\sigma^2}{n} M^{-1}(\mu^n)}(k_n + 2, m) = \frac{n}{\sigma^2} \mid\mid \widehat{\Theta} - \widehat{\Theta}_{m,k_n+2} \mid\mid^2_{M(\mu^n)} .$$

But, for m sufficiently large and for fixed n, (see section 3.2.3) we have: $\bar{\chi}^2_{\frac{\sigma^2}{n} M^{-1}(\mu^n)}(k_n + 2, m)$ converges in probability to

$$\bar{\chi}^2_{\frac{\sigma^2}{n} M^{-1}(\mu^n)}(k_n + 2) = \frac{n}{\sigma^2} \mid\mid \widehat{\Theta} - I\!\!P(\widehat{\Theta}) \mid\mid^2_{M(\mu^n)}$$

where $I\!\!P(\widehat{\Theta})$is the $M(\mu^n)-$projection of $\widehat{\Theta}$ onto the polyhedral closed convex cone

$$\mathcal{K} = \{x \in I\!\!R^{k_n+4} \ : b'_l x \le 0, \quad l = 0, ..., k_n + 1\}.$$

Otherwise,

$$\bar{\chi}^2_{\frac{\sigma^2}{n} M^{-1}(\mu^n)}(k_n + 2) = \frac{n}{\sigma^2} \mid\mid (\widehat{\Theta} - {}_s\widehat{\Theta}) + ({}_s\widehat{\Theta} - I\!\!P({}_s\widehat{\Theta})) - I\!\!P(\widehat{\Theta}) + I\!\!P({}_s\widehat{\Theta}) \mid\mid^2_{M(\mu^n)} .$$

We can rewritten this in the following form:

$$\bar{\chi}^2_{\frac{\sigma^2}{n}M^{-1}(\mu^n)}(k_n+2) = \frac{n}{\sigma^2} \mid\mid {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \mid\mid^2_{M(\mu^n)}$$

$$+\frac{n}{\sigma^2} \mid\mid B_n \mid\mid^2_{M(\mu^n)} + \frac{n}{\sigma^2} \mid\mid I\!P({}_s\widehat{\Theta}) - I\!P({}_s\widehat{\Theta} + B_n) \mid\mid^2_{M(\mu^n)}$$

$$+\frac{2n}{\sigma^2} < {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}), B_n >_{M(\mu^n)} + \frac{2n}{\sigma^2} < {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}), I\!P({}_s\widehat{\Theta}) - I\!P({}_s\widehat{\Theta} + B_n) >_{M(\mu^n)}$$

$$+\frac{2n}{\sigma^2} < B_n, I\!P({}_s\widehat{\Theta}) - I\!P({}_s\widehat{\Theta} + B_n) >_{M(\mu^n)}$$

where $<,>_{M(\mu^n)}$ is the saclar product defined by the metric $M(\mu^n)$.
Now, let us first show that, under the null hypothesis, we have:

$$\sup_{f\in\mathcal{F}_{l,M}} (\bar{\chi}^2_{\frac{\sigma^2}{n}M^{-1}(\mu^n)}(k_n+2) - \frac{n}{\sigma^2} \mid\mid {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \mid\mid^2_{M(\mu^n)})$$

converges in probability to zero. Indeed, the following inequality hold:

$$\sup_{f\in\mathcal{F}_{l,M}} \sqrt{\frac{n}{\sigma^2} \mid\mid B_n \mid\mid^2_{M(\mu^n)}} \leq \sup_{f\in\mathcal{F}_{l,M}} \sqrt{\frac{n}{\sigma^2} \mid\mid B_n \mid\mid^2 \mid\mid M(\mu^n) \mid\mid}$$

Now, it is easily seen, as in Diack&Thomas (1998) (see formula (3.5)) that the right-hand size of the above inequality verifie the following equality:

$$\sup_{f\in\mathcal{F}_{l,M}} \sqrt{\frac{n}{\sigma^2} \mid\mid B_n \mid\mid^2 \mid\mid M(\mu^n) \mid\mid} = \mathcal{O}(r^{1/2} n^{1/2} \delta_{k_n}{}^4 (\sup_{1\leq i\leq r} \mu_i)^{1/2}).$$

Hence,

$$\sup_{f\in\mathcal{F}_{l,M}} \sqrt{\frac{n}{\sigma^2} \mid\mid B_n \mid\mid^2_{M(\mu^n)}} = \mathcal{O}(r^{1/2} n^{1/2} \delta_{k_n}{}^4 (\sup_{1\leq i\leq r} \mu_i)^{1/2})$$

Therefore, using (iii) of theorem 5, we see that

$$\lim_{n\longrightarrow+\infty}\sup \sup_{f\in\mathcal{F}_{l,M}} \sqrt{\frac{n}{\sigma^2} \mid\mid B_n \mid\mid^2_{M(\mu^n)}} = 0. \tag{3.31}$$

Using now the fact that projections onto closed convex cones are contracting maps as are projections onto linear subspaces, we obtain that

$$\frac{n}{\sigma^2} \mid\mid I\!P({}_s\widehat{\Theta}) - I\!P({}_s\widehat{\Theta} + B_n) \mid\mid^2_{M(\mu^n)} \leq \frac{n}{\sigma^2} \mid\mid B_n \mid\mid^2_{M(\mu^n)} .$$

Then

$$\lim_{n\longrightarrow+\infty}\sup \sup_{f\in\mathcal{F}_{l,M}} \frac{n}{\sigma^2} \mid\mid I\!P({}_s\widehat{\Theta}) - I\!P({}_s\widehat{\Theta} + B_n) \mid\mid^2_{M(\mu^n)} = 0 \quad a.s.$$

It follows that

$$\lim_{n\longrightarrow+\infty}\sup \sup_{f\in\mathcal{F}_{l,M}} \frac{2n}{\sigma^2} < B_n, I\!P({}_s\widehat{\Theta}) - I\!P({}_s\widehat{\Theta} + B_n) >_{M(\mu^n)} = 0 \quad a.s.$$

On the other hand

$$\left| \frac{2n}{\sigma^2} < {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}), B_n >_{M(\mu^n)} \right| \leq \frac{2n}{\sigma^2} \| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \|_{M(\mu^n)} \| B_n \|_{M(\mu^n)}$$

$$\leq \frac{2n}{\sigma^2} \| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \|_{M(\mu^n)} \| B_n \|_{M(\mu^n)} \, 1\!\!I_{(\sqrt{n}\| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta})\|_{M(\mu^n)} \leq 1)}$$

$$+ \frac{2n}{\sigma^2} \| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \|_{M(\mu^n)} \| B_n \|_{M(\mu^n)} \, 1\!\!I_{(\sqrt{n}\| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta})\|_{M(\mu^n)} > 1)}$$

where $1\!\!I_A$ is defined by:

$$1\!\!I_A(x) = 1 \quad \text{if} \quad x \in A \quad \text{and} \quad 1\!\!I_A(x) = 0 \quad \text{otherwise.}$$

By (3.31), we have

$$\lim \sup_{n \longrightarrow +\infty} \sup_{f \in \mathcal{F}_{l,M}} \frac{2n}{\sigma^2} \| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \|_{M(\mu^n)} \| B_n \|_{M(\mu^n)} \, 1\!\!I_{(\sqrt{n}\| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta})\|_{M(\mu^n)} \leq 1)} = 0.$$

Otherwise,

$$\frac{2n}{\sigma^2} \| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \|_{M(\mu^n)} \| B_n \|_{M(\mu^n)} \, 1\!\!I_{(}\sqrt{n} \| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \|_{M(\mu^n)} > 1)$$

$$\leq \frac{2n}{\sigma^2} \| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \|^2_{M(\mu^n)} \sqrt{n} \| B_n \|_{M(\mu^n)} \, 1\!\!I_{(\sqrt{n}\| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta})\|_{M(\mu^n)} > 1)}$$

Now, under the null hypothesis, and for n sufficiently large, S is convex. That is to say that $\Theta \in \mathcal{K}$. Then, using lemma 2, we see that

$$\| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \|_{M(\mu^n)} \leq \| {}_s\widehat{\Theta} - \Theta - I\!P({}_s\widehat{\Theta} - \Theta) \|_{M(\mu^n)} .$$

Moreover,

$${}_s\widehat{\Theta} - \Theta \rightsquigarrow \mathcal{N}(0, \frac{\sigma^2}{n} M^{-1}(\mu^n))$$

Applying now, theorem 2, we obtain

$$P(\frac{n}{\sigma^2} \| {}_s\widehat{\Theta} - \Theta - I\!P({}_s\widehat{\Theta} - \Theta) \|^2_{M(\mu^n)} \geq s) = \omega_0 P(\chi^2_0 \geq s^2) + (\sum_{j=0}^{k_n+1} \omega_j) P(\chi^2_1 \geq s^2)$$

$$+ \omega_{k_n+4} P(\chi^2_{k_n+4} \geq s^2).$$

Then, under the null hypothesis,

$$P(\frac{n}{\sigma^2} \| {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \|^2_{M(\mu^n)} \geq s) \leq P(\chi^2_{k_n+4} \geq s^2).$$

Hence, for all $\epsilon \in I\!R^{+*}$, we have, on a

$$P(\sup_{f\in\mathcal{F}_{l,M}} \frac{2n}{\sigma^2} \mid\mid {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \mid\mid_{M(\mu^n)} \mid\mid B_n \mid\mid_{M(\mu^n)} \mathbb{I}_{(}\sqrt{n} \mid\mid {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \mid\mid_{M(\mu^n)} > 1) > \epsilon)$$

$$\leq P(\sup_{f\in\mathcal{F}_{l,M}} \chi^2_{k_n+4}\sqrt{n} \mid\mid B_n \mid\mid_{M(\mu^n)} > \epsilon).$$

Applying now the Bienayme-Chebychev inequality, we can deduce that

$$P(\sup_{f\in\mathcal{F}_{l,M}} \frac{2n}{\sigma^2} \mid\mid {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \mid\mid_{M(\mu^n)} \mid\mid B_n \mid\mid_{M(\mu^n)} \mathbb{I}_{(}\sqrt{n} \mid\mid {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \mid\mid_{M(\mu^n)} > 1) > \epsilon)$$

$$\leq \frac{n \mid\mid B_n \mid\mid^2_{M(\mu^n)}}{\epsilon^2} I\!E((\chi^2_{k_n+4})^2).$$

We have

$$I\!E((\chi^2_{k_n+4})^2) = 2(k_n+4) = \mathcal{O}(\delta_{k_n}).$$

Then

$$P(\sup_{f\in\mathcal{F}_{l,M}} \frac{2n}{\sigma^2} \mid\mid {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \mid\mid_{M(\mu^n)} \mid\mid B_n \mid\mid_{M(\mu^n)} \mathbb{I}_{(}\sqrt{n} \mid\mid {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \mid\mid_{M(\mu^n)} > 1) > \epsilon)$$

$$= \mathcal{O}(r^{1/2} n^{1/2} {\delta_{k_n}}^3 (\sup_{1\leq i\leq r} \mu_i)^{1/2}).$$

Hence under (iii) of theorem 5, we obtain

$$\lim \sup_{n\longrightarrow+\infty} \sup_{f\in\mathcal{F}_{l,M}} \mid \frac{2n}{\sigma^2} < {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}), B_n >_{M(\mu^n)} \mid = 0$$

with convergence in probability.
Now,

$$\mid \frac{2n}{\sigma^2} < {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}), I\!P({}_s\widehat{\Theta}) - I\!P({}_s\widehat{\Theta} + B_n) >_{M(\mu^n)} \mid \leq \frac{2n}{\sigma^2} \mid\mid {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \mid\mid_{M(\mu^n)} \mid\mid B_n \mid\mid_{M(\mu^n)}.$$

We see as above that

$$\sup_{f\in\mathcal{F}_{l,M}} \mid \frac{2n}{\sigma^2} < {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}), I\!P({}_s\widehat{\Theta}) - I\!P({}_s\widehat{\Theta} + B_n) >_{M(\mu^n)} \mid$$

converges in probability to zero. Hence

$$\sup_{f\in\mathcal{F}_{l,M}} (\bar{\chi}^2_{\frac{\sigma^2}{n}M^{-1}(\mu^n)}(k_n+2) - \frac{n}{\sigma^2} \mid\mid {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \mid\mid^2_{M(\mu^n)})$$

converges in probability to zero. Therefore, we can write that

$$\lim \sup_{n\longrightarrow+\infty} \lim_{m\longrightarrow+\infty} \sup_{f\in H_0} \mathcal{P}_f(\Lambda_{n,m}) = \lim \sup_{n\longrightarrow+\infty} \sup_{f\in H_0} \mathcal{P}_f(\bar{\chi}^2_{\frac{\sigma^2}{n}M^{-1}(\mu^n)}(k_n+2) \geq s^2_{\alpha,k_n+2})$$

$$= \lim \sup_{n \longrightarrow +\infty} \sup_{f \in H_0} \mathcal{P}_f\left(\frac{n}{\sigma^2} \mid\mid {}_s\widehat{\Theta} - I\!P({}_s\widehat{\Theta}) \mid\mid^2_{M(\mu^n)} \geq s^2_{\alpha,k_n+2}\right)$$

$$\leq \lim \sup_{n \longrightarrow +\infty} \sup_{f \in H_0} \mathcal{P}_f\left(\frac{n}{\sigma^2} \mid\mid {}_s\widehat{\Theta} - \Theta - I\!P({}_s\widehat{\Theta} - \Theta) \mid\mid^2_{M(\mu^n)} \geq s^2_{\alpha,k_n+2}\right) = \alpha.$$

Otherwise $f \equiv 0$ lies in $H_\circ$ and in this point, we have

$$\lim \sup_{n \longrightarrow +\infty} \lim_{m \longrightarrow +\infty} \mathcal{P}_f(\Lambda_{n,m}) = \alpha.$$

Then the test is asymptotically of size α. It remainds to prove that the test is consistent. From theorem 3, it suffices to show that

$$\frac{n}{\sigma^2} \mid\mid I\!E\widehat{\Theta} - I\!P(I\!E\widehat{\Theta}) \mid\mid^2_{M(\mu^n)} \longrightarrow +\infty.$$

As above, we have

$$\frac{n}{\sigma^2} \mid\mid I\!E\widehat{\Theta} - I\!P(I\!E\widehat{\Theta}) \mid\mid^2_{M(\mu^n)} = \frac{n}{\sigma^2} \mid\mid B_n + \Theta - I\!P(B_n + \Theta) \mid\mid^2_{M(\mu^n)} .$$

In other words

$$\frac{n}{\sigma^2} \mid\mid I\!E\widehat{\Theta} - I\!P(I\!E\widehat{\Theta}) \mid\mid^2_{M(\mu^n)} = \frac{n}{\sigma^2} \mid\mid \Theta - I\!P(\Theta) \mid\mid^2_{M(\mu^n)} + \frac{n}{\sigma^2} \mid\mid B_n \mid\mid^2_{M(\mu^n)}$$

$$+ \frac{n}{\sigma^2} \mid\mid I\!P(\Theta) - I\!P(\Theta + B_n) \mid\mid^2_{M(\mu^n)} + \frac{2n}{\sigma^2} < \Theta - I\!P(\Theta), B_n >_{M(\mu^n)}$$

$$+ \frac{2n}{\sigma^2} < \Theta - I\!P(\Theta), I\!P(\Theta) - I\!P(\Theta + B_n) >_{M(\mu^n)} + \frac{2n}{\sigma^2} < B_n, I\!P(\Theta) - I\!P(\Theta + B_n) >_{M(\mu^n)} .$$

In the same way as above, we see that

$$\frac{n}{\sigma^2} \mid\mid B_n \mid\mid^2_{M(\mu^n)} \to 0.$$

$$\frac{n}{\sigma^2} \mid\mid I\!P(\Theta) - I\!P(\Theta + B_n) \mid\mid^2_{M(\mu^n)} \to 0.$$

Hence

$$\frac{2n}{\sigma^2} \left| < \Theta - I\!P(\Theta), B_n >_{M(\mu^n)} \right| = \epsilon_n\left(\frac{n}{\sigma^2} \mid\mid \Theta - I\!P(\Theta) \mid\mid^2_{M(\mu^n)}\right)$$

and

$$\frac{2n}{\sigma^2} \left| < \Theta - I\!P(\Theta), I\!P(\Theta) - I\!P(\Theta + B_n) >_{M(\mu^n)} + \frac{2n}{\sigma^2} < B_n, I\!P(\Theta) - I\!P(\Theta + B_n) >_{M(\mu^n)} \right|$$

$$= \epsilon'_n\left(\frac{n}{\sigma^2} \mid\mid \Theta - I\!P(\Theta) \mid\mid^2_{M(\mu^n)}\right)$$

where ϵ_n and ϵ'_n are nonnegatives reals converging to zero. Therefore, the consistency of the test will be established when we show that

$$\frac{n}{\sigma^2} \mid\mid \Theta - I\!P(\Theta) \mid\mid^2_{M(\mu^n)} \longrightarrow +\infty.$$

If f is non-convex, then S is also non-convex and therefore, $\Theta \notin \mathcal{K}$. Hence, there is a face F_j of $\mathcal{K}$ such that $I\!P(\Theta)$ lies in F_j with F_j defined by

$$F_j = \{x \in I\!R^{k_n+4} \ : b_j' x = 0, \quad b_l' x \leq 0, \quad l = 0, ..., k_n + 1\}.$$

Hence, from lemma 1, $I\!P(\Theta)$ is also the orthogonal projection of Θ onto the subspace generated by F_j and $\Theta - I\!P(\Theta)$ is also in the polar cone of $\mathcal{K}$. That is to say

$$\frac{n}{\sigma^2} \mid\mid \Theta - I\!P(\Theta) \mid\mid^2_{M(\mu^n)} = \frac{n(b'_j\Theta)^2}{\sigma^2(b'_j M^{-1}(\mu^n) b_j)}.$$

And

$$\begin{cases} b'_j\Theta \leq 0 \\ b_i'\Theta - \frac{b'_i M^{-1}(\mu^n) b_j}{b'_j M^{-1}(\mu^n) b_j} b_j'\Theta \leq 0, \quad i = 0, ..., k_n + 1. \end{cases}$$

Otherwise,

$$\frac{1}{b'_j M^{-1}(\mu^n) b_j} \geq \frac{1}{\mid\mid b_j \mid\mid^2 \mid\mid M^{-1}(\mu^n) \mid\mid}.$$

Hence,

$$\frac{n}{\sigma^2} \mid\mid \Theta - I\!P(\Theta) \mid\mid^2_{M(\mu^n)} \geq \frac{n(b'_j\Theta)^2}{\sigma^2 \mid\mid b_j \mid\mid^2 \mid\mid M^{-1}(\mu^n) \mid\mid}.$$

Therefore, for n sufficiently large,

$$\frac{n}{\sigma^2} \mid\mid \Theta - I\!P_2(\Theta) \mid\mid^2_{M(\mu^n)} \geq \frac{n}{\sigma^2}(f''(\eta_j))^2 \delta^5_{k_n}.$$

Under the assumption (iv) on a $n\delta^5_{k_n} \to +\infty$ hence, we will finished when we show that: for n sufficiently large $(f''(\eta_j))^2 > \epsilon$ where ϵ is a positif real.
Let $a_{ij} = \frac{b'_i M^{-1}(\mu^n) b_j}{b'_j M^{-1}(\mu^n) b_j}$. Because the sequence of knots is quasi-uniforme, it is readily to see that

$$\sup_{0 \leq j \leq k_n+1} \mid\mid b_j \mid\mid = \mathcal{O}(\inf_{0 \leq j \leq k_n+1} \mid\mid b_j \mid\mid).$$

Since now,

$$\frac{1}{b'_j M^{-1}(\mu^n) b_j} \leq \frac{\mid\mid M(\mu^n) \mid\mid}{\mid\mid b_j \mid\mid^2} \leq \frac{\mid\mid M(\mu^n) \mid\mid}{\inf_{0 \leq j \leq k_n+1} \mid\mid b_j \mid\mid^2}.$$

We see that (using the fact that $\mid\mid M(\mu^n) \mid\mid = \mathcal{O}(\delta_{k_n})$ and $\mid\mid M^{-1}(\mu^n) \mid\mid = \mathcal{O}(\delta^{-1}_{k_n})$ (see Diack& Thomas (1998)))

$$\mid a_{ij} \mid \leq \frac{\sup_{0 \leq j \leq k_n+1} \mid\mid b_j \mid\mid^2 \mid\mid M^{-1}(\mu^n) \mid\mid \mid\mid M(\mu^n) \mid\mid}{\inf_{0 \leq j \leq k_n+1} \mid\mid b_j \mid\mid^2} = \mathcal{O}(1).$$

Now, let us suppose That $b_i'\Theta \geq 0$ since $b_j'\Theta \leq 0$ we have $a_{ij} \leq 0$. Then

$$0 \leq b_i'\Theta \leq a_{ij} b_j'\Theta \leq -\mathcal{O}(1) b_j'\Theta.$$

Hence,

$$(b_i^{'}\Theta)^2 \leq \mathcal{O}(1)(b_j^{'}\Theta)^2.$$

But for n sufficiently large,

$$b_i^{'}\Theta = -f''(\eta_i) + \circ(1).$$

Because the knots are dense in $(0,1)$ one can deduce that

$$(b_j^{'}\Theta)^2 = (f''(\eta_j) + \circ(1))^2 \leq c \sup_{x\in(0,1)} [(f''(x))^2 \mathbb{I}_{(f''(x)\leq 0)}] > \epsilon$$

where c et ϵ are positive constant. We have

$$\frac{n}{\sigma^2} \mid\mid \Theta - I\!P(\Theta) \mid\mid^2_{M(\mu^n)} \rightarrow +\infty.$$

The consistency of the test follows OUf !□

References

[Agarwal and Studden, 1980] Agarwal, G. and Studden, W. (1980). Asymptotic integrated mean square error using least squares and bias minimizing splines. *The Annals of Statistics*, 8(6):1307–1325.

[Barlow et al., 1972] Barlow, R. E., Bartholomew, D., Bremner, J., and Brunck, H. (1972). *Statistical inference under order restrictions.* John Wiley α Sons Ltd.

[Bartholomew, 1961] Bartholomew, D. (1961). Ordered tests in the analysis of variance. *Biometrika*, 48:325–332.

[Beatson, 1982] Beatson, R. (1982). Monotone and convex approximation by splines: error estimates and a curve fitting algorithm. *SIAM J. of Math. Analysis*, 19(4).

[Carroll and Kimball, 1996] Carroll, C. and Kimball, M. (1996). Notes and comments on the concavity of the comsumption function. *Econometrica*, 64(4):981–992.

[Diack, 1996] Diack, C. (1996). *Testing convexity.* Proceedings in Computational Statistics: Physica Verlag.

[Diack, 1997] Diack, C. (1997). *Tests de convexite pour une fonction de regression.* These de Doctorat de l'Universite Paul Sabatier de Toulouse, France.

[Diack, 1998] Diack, C. (1998). Sur la convergence des tests de schlee et yatchew. *submitted.*

[Diack and Thomas-Agnan, 1998] Diack, C. and Thomas-Agnan, C. (1998). A nonparametric test of the non-convexity of regression. *to appear in Journal of Nonparametric Statistics*.

[Dierckx, 1980] Dierckx, H. (1980). An algorithm for cubic spline fitting with convexity constraints. *Computing*, 24:349–371.

[Dykstra, 1983] Dykstra, R. (December 1983). An algorithm for restricted least squares regression. *Journal of american statistical association*, 78(384).

[Kudo, 1963] Kudo, A. (1963). A multivariate analogue of the one-sided test. *Biometrika*, 50(3 and 4):403–418.

[Kudo and Choi, 1975] Kudo, A. and Choi, J. (1975). A generalized multivariate analogue of the one-sided test. *Mem. Fac. Sci., Kyushu Univ. A.*, 29:303–328.

[Nuesch, 1966] Nuesch, P. (1966). On the problem of testing location in multivariate populations for restricted alternatives. *Ann. Math. Statist.*, 37:113–119.

[Raubertas et al., 1986] Raubertas, R., Lee, C., and Nordheim, E. (1986). Hypothesis tests for normal means constrained by linear inequalities. *Comm. Statist. Theor. Meth.*, 15:2809–2833.

[Robertson et al., 1988] Robertson, T., Wright, F., and Dykstra, R. (1988). *Order restricted statistical inference*. Wiley New-York.

[Rockafeller, 1970] Rockafeller, R. (1970). *Convex Analysis*. Princeton, New Jersey: Princeton University Press.

[Sasabuchi, 1980] Sasabuchi, S. (1980). A test of a multivariate normal mean with composite hypotheses determined by linear inequalities. *Biometrika*, 67(2):429–439.

[Schlee, 1980] Schlee, W. (1980). Nonparametric test of the monotony and convexity of regression. *Nonparametric Statistical Inference*, 2:823–836.

[Schumaker, 1981] Schumaker, L. (1981). *Spline function: Basic theory*. John Wiley, New York.

[Shapiro, 1985] Shapiro, A. (1985). Asymptotic distribution of test statistics in the analysis of moment structures under inequality constraits. *Biometrika*, 72(1):133–144.

[Yatchew, 1992] Yatchew, A. (1992). Nonparametric regression tests based on least squares. *Econometrics Theory*, 8:435–451.